UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/626,668	07/25/2003	Terrence B. Peace	P23789.P06	1823

7055 7590 10/27/2006

GREENBLUM & BERNSTEIN, P.L.C.
1950 ROLAND CLARKE PLACE
RESTON, VA 20191

EXAMINER	
COUGHLAN, PETER D	

ART UNIT	PAPER NUMBER
2129	

DATE MAILED: 10/27/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/626,668	PEACE, TERRENCE B.	
	Examiner	Art Unit	
	Peter Coughlan	2129	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *25 August 2006*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-15 and 17-34* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-15 and 17-34* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *7/25/2003* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

Detailed Action

1. This office action is in response to an AMENDMENT entered August 25, 2006 for the patent application 10/626668 filed on May 3, 2002.

2. The First Office Action of July 25, 2003 is fully incorporated into this Final Office Action by reference.

Status of Claims

3. Claims 1-15, 17-34 are pending.

35 USC § 101

4. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1-15, 17-34 are rejected under 35 U.S.C. 101 for nonstatutory subject matter. The computer system must set forth a practical application of that § 101 judicial exception to produce a real-world result. Benson, 409 U.S. at 71-72, 175 USPQ at 676-77. The invention is ineligible because it has not been limited to a substantial practical

application. A method for testing the validity of a prediction model has no real world purpose. The result is a Boolean value or a numerical value that determines classification or not. 'Classification' itself is vague and indeterminate thus lacking a practical application.

Another flaw with the invention is it lacks concreteness. Data that is to be tested is compared to random generated data of a given distribution to determine if it matches. The problem with the invention is the fact the randomness of the determining requirements is random! It cannot be repeated and therefore lacks concreteness.

In determining whether the claim is for a "practical application," the focus is not on whether the steps taken to achieve a particular result are useful, tangible and concrete, but rather that the final result achieved by the claimed invention is "useful, tangible and concrete." If the claim is directed to a practical application of the § 101 judicial exception producing a result tied to the physical world that does not preempt the judicial exception, then the claim meets the statutory requirement of 35 U.S.C. § 101.

Comparing incoming data to random generated data has no function. Analyzing data has no purpose. A computer program that calculates, compares and determines data has no real world purpose.

Comparing real data to randomly generated data that is 'suppose to be of the same distribution' (One does not know for sure due to the fact it is randomly generated.) is not reliable due to the inherent nature one side of the comparison is randomly generated.

The invention must be for a practical application and either:

1) specify transforming (physical thing) or

2) have the FINAL RESULT (not the steps) achieve or produce a

useful (specific, substantial, AND credible),

concrete (substantially repeatable/ non-unpredictable), AND

tangible (real world/ non-abstract) result.

A claim that is so broad that it reads on both statutory and non-statutory subject matter, must be amended, and if the specification discloses a practical application but the claim is broader than the disclosure such that it does not require the practical application, then the claim must be amended.

Claims that recites a computer output that is of a numeric or Boolean value has no real world function and is not statutory. The claims need to be narrowed down to a practical application to overcome the 35 USC § 101 rejection.

Claims when following an algorithm and a portion of that algorithm is randomly generated for comparison lacks concreteness and therefore is not statutory. Applicant makes the argument of a large sample set that produces a non-random effect. Examiner disagrees and suggests a large sample set only increases a probability of a repeatable distribution probability but does not guarantee it.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1, 2, 3, 4, 5, 6, 7, 15, 17, 18, 20, 21, 26, 28, 29, 30, 31, 34 are rejected under 35 U.S.C. 103(a) as being unpatentable over Tatsuoka in view of Zhang (U. S. Patent 5855011, referred to as **Tatsuoka**; U. S. Patent 5832182, referred to as **Zhang**)

Claim 1.

Tatsuoka teaches specifying a test statistic formula (**Tatsuoka,** abstract); computing a numerical value NTS of the test statistic using the test statistic formula and the original data set (**Tatsuko,** C8:17-29; 'NTS' of applicant is equivalent to 'quality measure' of Tatsuko.); specifying a probability distribution relating to the original data set. (**Tatsuoka,** C15:11-21; 'Specifying' of applicant is equivalent to 'jumping' of Tatsuka.)

Tatsuoka does not teach creating a plurality of random data sets RDB(i) using randomly generated data, in which i is a positive integer.

Zhang teaches creating a plurality of random data sets RDB(i) using randomly generated data, in which i is a positive integer. (**Zhang,** C20:40-59; 'Data sets' of applicant is equivalent to 'each intended cluster' of Zhang.) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify the teachings of Tatsuoka by incrementing a vector into positive increments as taught by Zhang to create a plurality of random data sets RDB(i) using randomly generated data, in which i is a positive integer.

For the purpose of containing the data into a usable form for future analysis.

Tatsuoka teaches computing a plurality of numerical values TS(i) of the test statistic corresponding to the plurality of random data sets RDB(i), and storing each numerical value TS(i) in a numerical test statistic array (**Tatsuoka,** C16:1-8; 'TS(i)' of applicant is equivalent to 'test item counter' of Tatsuoka.); and comparing the numerical value NTS with the numerical test statistic array to determine a non-empty set of percentile values corresponding to the numerical value NTS and an associated non-empty set of percentile indices. (**Tatsuoka,** C16:9-17; 'Comparing' of applicant is equivalent to 'test subject classified' in step 19 of Tatsuoka.)

and determining whether to accept the prediction model based on the comparison of the numerical value NTS with the numerical test statistic array. (**Tatsuoka,** abstract and C2:46-51; 'Numerical value NTS' of applicant is equivalent to 'quality measure' of Tatsuoka. 'Numerical test statistic array' of applicant is equivalent to 'SPS' of Tatsuoka.)

Claim 2.

Tatsuoka does not teach each of the plurality of data sets RDB(i) is distributed according to the probability distribution.

Zhang teaches each of the plurality of data sets RDB(i) is distributed according to the probability distribution. (**Zhang,** C20:40-59; 'Probability distribution' of applicant is equivalent to 'distributed uniformly' of Zhang.) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify the teachings of Tatsuoka by randomly generating a probability distribution to a specific model as taught by Zhang to have a plurality of data sets RDB(i) is distributed according to the probability distribution.

For the purpose of comparing a given set of data to a randomly generated set of data.

Claims 3, 15, 28 and 29

Tatsuoka teaches in which each the data sets RDB(i) has a size that is functionally equivalent to a size of the original data set. (**Tatsuoka,** C1:43-54; 'Size of RDB(i)', 'equivalent' and 'original data set' of applicant is equivalent to 'test items', 'allowing' and 'model' of Tatsuoka.)

Claim 4.

Tatsuoka does not teach determining a null hypothesis defining a potential relationship among data in the original data set and rejecting the null hypothesis as not

accurately representing the original data set when the value of a function of the non-empty set of percentile indices, associated with the non-empty set of percentile values, which correspond to the numerical value NTS, is in an extreme range, indicating that the numerical value NTS did not arise by chance.

Zhang teaches determining a null hypothesis defining a potential relationship among data in the original data set (**Zhang,** C15:23-50; 'Determining a null hypothesis' of applicant is equivalent to 'new tree starts with null' of Zhang.); and rejecting the null hypothesis as not accurately representing the original data set when the value of a function of the non-empty set of percentile indices, associated with the non-empty set of percentile values, which correspond to the numerical value NTS, is in an extreme range, indicating that the numerical value NTS did not arise by chance. (**Zhang,** C15:23 through C16:15; Rejection of the null hypothesis is done by the function 'Re-build'. The numerical value of 'NTS' of applicant is equivalent to the value generated by 'ClosestPath' of Zhang.) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify the teachings of Tatsuoka by comparing a value of the incoming data to a value of the randomly generated data as taught by Zhang to determine a null hypothesis defining a potential relationship among data in the original data set and rejecting the null hypothesis as not accurately representing the original data set when the value of a function of the non-empty set of percentile indices, associated with the non-empty set of percentile values, which correspond to the numerical value NTS, is in an extreme range, indicating that the numerical value NTS did not arise by chance.

For the purpose of determining is a correlation exists.

Claims 5 and 18.

Tatsuoka does not teach in which the non-empty set of percentile values comprises the greatest percentile value less than NTS and the smallest percentile value greater than NTS, and the non-empty set of percentile indices comprises the two percentile indices corresponding to the two percentile values of the non-empty set of percentile values.

Zhang teaches in which the non-empty set of percentile values comprises the greatest percentile value less than NTS and the smallest percentile value greater than NTS, and the non-empty set of percentile indices comprises the two percentile indices corresponding to the two percentile values of the non-empty set of percentile values. (**Zhang,** C15:23 through C16:15; 'Smallest percentile' of applicant is equivalent to 'CurentPath' of Zhang. 'Greatest percentile' of applicant is accomplished by the Boolean function 'status' of Zhang.) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify the teachings of Tatsuoka by using the highest and lowest percentile values as taught by Zhang to have both the non-empty set of percentile values comprises the greatest percentile value less than NTS and the smallest percentile value greater than NTS, and the non-empty set of percentile indices comprises the two percentile indices corresponding to the two percentile values of the non-empty set of percentile values.

For the purpose of using the greatest difference between two data sets for calculation of determining a given distribution.

Claim 6.

Tatsuoka teaches in which one percentile index is selected, when the corresponding percentile value meets a predetermined criterion for proximity to the numerical value NTS of the test statistic corresponding to the original data set. (**Tatsuoka,** C16:9-17; 'Proximity' of applicant is equivalent to 'threshold' of Tatsuoka.)

Claims 7 and 21.

Tatsuoka teaches in which the function of percentile indices is a linear combination of the non-empty set of percentile indices. (**Tatsuoka,** C6:51 through C7:4; Tatsuoka illustrates a linear function using summation of percentile indices.)

Claims 17 and 26.

Tatsuoka does not teach determining a plurality of percentile values, based on the plurality of numerical values, and a plurality of percentile indices corresponding to the plurality of percentile values ; and determining a non-empty set of selected percentile indices from the plurality of percentile indices, corresponding to the plurality of random data sets, by determining a non-empty set of percentile values from the plurality of percentile values which meets a predetermined criterion for proximity to the numerical value of the test statistic corresponding to the original data set.

Zhang teaches determining a plurality of percentile values, based on the plurality of numerical values, and a plurality of percentile indices corresponding to the plurality of percentile values (**Zhang,** C15:23 through C16:15; A plurality of numerical values of applicant is equivalent to 'CurentPath' and the Boolean function 'status' of Zhang.); and determining a non-empty set of selected percentile indices from the plurality of percentile indices, corresponding to the plurality of random data sets, by determining a non-empty set of percentile values from the plurality of percentile values which meets a predetermined criterion for proximity to the numerical value of the test statistic corresponding to the original data set. (**Zhang,** C15:23 through C16:15; The generation of the values for 'CurrentPath' and the Boolean function 'status' are predetermined.) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify the teachings of Tatsuoka by generating values based on analysis and determining if the relationship between the values classifies given a proximity model as taught by Zhang to determine a plurality of percentile values, based on the plurality of numerical values, and a plurality of percentile indices corresponding to the plurality of percentile values ; and determining a non-empty set of selected percentile indices from the plurality of percentile indices, corresponding to the plurality of random data sets, by determining a non-empty set of percentile values from the plurality of percentile values which meets a predetermined criterion for proximity to the numerical value of the test statistic corresponding to the original data set.

For the purpose of determining if given set of data matches a known distribution.

Claim 20.

Tatsuoka teaches the computer executable code further causing the computing device to determine that the numerical value of the test statistic corresponding to the original data set did not arise by chance when the value of a predetermined function of the selected percentile indices is outside a predetermined range of the plurality of percentile indices indicating numerical values that did arise by chance. (**Tatsuoka,** abstract; 'Not arise on chance' of applicant is equivalent to 'decision-theoretic rules' of Tatsuoka.)

Claim 30.

Tatsuoka teaches the program further comprising: a distribution determining source code segment that determines the distribution of the original data set by comparing the original data set with a plurality of theoretical distributions. (**Tatsuoka,** C23:59 through C24:22)

Claim 31.

Tatsuoka teaches a distribution determining source code segment that determines the distribution of the original data set by sorting the data into bins along at least one dimension. (**Tatsuoka,** C23:59 through C24:22; Tatsuoka uses Bayes theorem which calculates all possible probabilities of all combinations of a situation. For example if a situation has 3 outcomes A, B and C with respective probabilities of 0.2, 0.3 and 0.5 then Tatsuoka illustrates 3 bins 'A', 'B' and 'C'.)

Claim 34.

Tatsuoka teaches the program further comprising a distribution determining source code segment that determines an empirical distribution of the original data set. (**Tatsuoka,** C23:59 through C24:22; Tatsuoka illustrates a decision rule that analyzes a posterior distribution.)

Claim Rejections - 35 USC § 103

Claims 8, 9, 10, 11, 19, 22, 23, 27, 33 are rejected under 35 U.S.C. 103(a) as being unpatentable over the combination of Tatsuoka, and Zhang, as set forth above, and further in view of Shen (U. S. Patent 6041788, referred to as **Shen**)

Claim 8.

Tatsuoka, and Zhang do not teach in which the test statistic comprises a function of prediction error.

Shen teaches in which the test statistic comprises a function of prediction error. (**Shen,** C3:1-28; 'Prediction error' of applicant is equivalent to 'confidence interval' of Shen.) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by using a function that determines prediction error as taught by Shen to have the test statistic comprises a function of prediction error.

For the purpose of determining if the results fall within a given range.

Claim 9.

Tatsuoka, and Zhang do not teach in which the extreme range comprises one of above a 97.5.sup.th percentile and below a 2.5.sup.th percentile.

Shen teaches in which the extreme range comprises one of above a 97.5.sup.th percentile and below a 2.5.sup.th percentile. (**Shen,** C3:1-28; Above 97.5% or below 2.5% is a range of 95%, +/- 47.5%) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by giving a specific range for a prediction of error as taught by Shen to have the extreme range comprises one of above a 97.5.sup.th percentile and below a 2.5.sup.th percentile.

For the purpose of using 95% verses 5% for determination of an extreme range.

Claims 10 and 23.

Tatsuoka, and Zhang do not teach in which creating the plurality of random data sets RDB(i) comprises using randomly generated data according to a Monte Carlo technique.

Shen teaches in which creating the plurality of random data sets RDB(i) comprises using randomly generated data according to a Monte Carlo technique. (**Shen,** C2:54-67) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and

Zhang by using Monte Carlo for generating random numbers as taught by Shen to create the plurality of random data sets RDB(i) comprises using randomly generated data according to a Monte Carlo technique.

For the purpose of using a tested algorithm for the generation of random numbers.

Claim 11.

Tatsuoka, and Zhang do not teach constructing a confidence interval for the test statistic.

Shen teaches constructing a confidence interval for the test statistic. (**Shen,** C3:1-28; Shen illustrates a confidence interval above 95%) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by setting the boundaries for what is classified and what is not classified as taught by Shen to construct a confidence interval for the test statistic.

For the purpose of setting definite limits and these limits can be changed at a later dare in needed.

Claim 19

Tatsuoka, and Zhang do not teach the computer executable code further causing the computing device to select one percentile index when the corresponding percentile

value meets a predetermined criterion for proximity to the numerical value of the test statistic corresponding to the original data set.

Shen teaches the computer executable code further causing the computing device to select one percentile index when the corresponding percentile value meets a predetermined criterion for proximity to the numerical value of the test statistic corresponding to the original data set. (**Shen,** C3:1-28; Shen illustrates a 95% range with selection occurring when outside the 95% distribution.) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by having the method select a value index when a threshold has been exceeded as taught by Shen to have the computer executable code further causing the computing device to select one percentile index when the corresponding percentile value meets a predetermined criterion for proximity to the numerical value of the test statistic corresponding to the original data set.

For the purpose of using the computer to flag an instance when it has exceeded a given threshold limit.

Claim 22.

Tatsuoka, and Zhang do not teach in which the computer executable code further causes the computing device to construct a confidence interval for the test statistic.

Shen teaches in which the computer executable code further causes the computing device to construct a confidence interval for the test statistic. (**Shen,** C3:1-28; 'Prediction error' of applicant is equivalent to 'confidence interval' of Shen.) It would

have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by using the computer to generate threshold limits as taught by Shen have the computer executable code further causes the computing device to construct a confidence interval for the test statistic.

For the purpose of using a function within the code to adjust the threshold limits as needed based on a number of input parameters.

Claim 27.

Tatsuoka, and Zhang do not teach in which the range of values is based on the plurality of associated percentile indices.

Shen teaches in which the range of values is based on the plurality of associated percentile indices. (**Shen,** C3:1-28; Shen illustrates a generation of a distribution which is a range of values.) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by using input parameters to determine threshold limits as taught by Shen to have the range of values is based on the plurality of associated percentile indices.

For the purpose of taking into account the size, dimensions and distributions of the inputted data.

Claim 33.

Tatsuoka, and Zhang do not teach a confidence interval source code segment that constructs a confidence interval for the test statistic.

Shen teaches a confidence interval source code segment that constructs a confidence interval for the test statistic. (**Shen,** C3:1-28; Shen illustrates the generation of a confidence interval.) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by having the code generate a region of confidence as taught by Shen to have a confidence interval source code segment that constructs a confidence interval for the test statistic.

For the purpose of using this information to determine classification.

Claim Rejections - 35 USC § 103

Claims 12, 13, 14, 24, 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over the combination of Tatsuoka, and Zhang, as set forth above, and further in view of Fayyad. (U. S. Patent 6012058, referred to as **Fayyad**)

Claims 12, 14 and 25.

Tatsuoka, and Zhang do not teach in which each of the plurality of data sets RDB(i) has the same size, dimension and distribution as the original data set.

Fayyad teaches in which each of the plurality of data sets RDB(i) has the same size (**Fayyad,** C5:7-26; 'Size' of applicant is equivalent to 'M' of Fayyad.), dimension

(**Fayyad,** C1:30-37) and distribution (**Fayyad,** C2:5-15) as the original data set. It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by taking intro account the size, dimension and distribution of data as taught by Fayyad to have in which each of the plurality of data sets RDB(i) has the same size, dimension and distribution as the original data set.

For the purpose of flexibility so the method can render an refined answer due to the fact the results are based upon the input characteristics.

Claim 13.

Tatsuoka teaches a computing device for executing computer readable code. (**Tatsuoka,** C1:6-10; If Tatsuoka is computer implemented then it can read and execute computer code.)

Tatsuoka does not teach an input device for receiving data, the input device being in communication with the computing device.

Zhang teaches an input device for receiving data, the input device being in communication with the computing device. (**Zhang,** abstract) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by using an input device as taught by Fayyad to have an input device for receiving data, the input device being in communication with the computing device.

For the purpose of having the ability to input the device so the method can operate.

Tatsuoka teaches at least one data storage device for storing computer data, the data storage device being in communication with the computing device (**Tatsuoka,** C7:27-30); and a programming code reading device that reads computer executable code, the programming code reading device being in communication with the computing device. (**Tatsuoka,** C1:6-10 If Tatsuoka is computer implemented means it can read and execute computed code, this implies there is a device that can read the code. Computer keyboard, hard drive, floppy disk, disk, or flash drive are all components of 'computer implemented'.)

Tatsuoka and Zhang do not teach the computer executable code causing the computing device to generate a plurality of random data sets, each random data set having a second size, dimension and distribution relating to the original data set.

Fayyad teaches the computer executable code causing the computing device to generate a plurality of random data sets, each random data set having a second size (**Fayyad,** C5:7-26; 'Size' of applicant is equivalent to 'M' of Fayyad.), dimension (**Fayyad,** C1:30-37) and distribution (**Fayyad,** C2:5-15) relating to the original data set. It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by using parameters as size, dimension and distribution as taught by Fayyad to have the computer executable code causing the computing device to generate a plurality of

random data sets, each random data set having a second size, dimension and distribution relating to the original data set.

For the purpose of flexibility so the method can render an refined answer due to the fact the results are based upon the input characteristics.

Tatsuoka teaches calculate a plurality of numerical values of test statistics corresponding to the plurality of random data sets, each numerical value being calculated according to a test statistic formula (**Tatsuko,** C8:17-29; 'NTS' of applicant is equivalent to 'quality measure' of Tatsuko.); and to compare the plurality of numerical values and the numerical value of the test statistic corresponding to the original data set, calculated in accordance with the test statistic formula. (**Tatsuoka,** C16:9-17; 'Determining' of applicant is equivalent to 'test subject classified' in step 19 of Tatsuoka.)

wherein a determination is made, based on the comparison between the plurality of numerical values and the numerical value corresponding to the original data set, whether the original data set is characterized by at least one factor that is not based on chance. (**Tatsuoka,** abstract; 'Not based on chance' of applicant is equivalent to 'decision-theoretic rules' of Tatsuoka. 'Numerical measure' of applicant is equivalent to 'quality measure' of Tatsuoka.)

Claim 24.

Tatsuoka teaches a calculating source code segment that calculates a plurality of numerical values of test statistics corresponding to a plurality of randomly generated

data sets, calculated in accordance with the predetermined test statistic formula. (**Tatsuoka,** C16:1-8)

Tatsuoka and Zhang do not teach each randomly generated data set having a second size, dimension and distribution relating to the original data set.

Fadday teaches each randomly generated data set having a second size (**Fayyad,** C5:7-26; 'Size' of applicant is equivalent to 'M' of Fayyad.), dimension (**Fayyad,** C1:30-37) and distribution relating to the original data set. (**Fayyad,** C2:5-15) It would have been obvious to a person having ordinary skill in the art at the time of applicant's invention to modify combined teachings of Tatsuoka, and Zhang by using input parameters such as size, dimension, and distribution as taught by Fayyad tohave each randomly generated data set having a second size, dimension and distribution relating to the original data set.

For the purpose of flexibility so the method can render an refined answer due to the fact the results are based upon the input characteristics.

Tatsuoka teaches a comparing source code segment that compares a numerical value of a test statistic calculated in accordance with the predetermined test statistic formula and calculated with the original data set, with the plurality of numerical values corresponding to the plurality of randomly generated data sets (**Tatsuoka,** C16:9-17; 'Comparing' of applicant is equivalent to 'test subject classified' in step 19 of Tatsuoka.); and a determining source code segment that determines that at least one factor in the original data set did not arise by chance when the numerical value of the test statistic calculated from the original data set is not within a range, within the plurality of

numerical values corresponding to the plurality of randomly generated data sets, representative of numerical values arising by chance; (**Tatsuoka,** abstract; 'Not arise on chance' of applicant is equivalent to 'decision-theoretic rules' of Tatsuoka.)

wherein a hypothesis is accepted when the original data set is determined as not arising by chance. (**Tatsuoka,** abstract; 'Not arising by chance' of applicant is equivalent to 'decision-theoretic rules' of Tatsuoka.)

Response to Arguments

5. Applicant's arguments filed on August 25, 2006 for claims 1-15, 17-34 have been fully considered but are not persuasive.

6. In reference to the Applicant's argument:

Initially, Applicant would like to express his appreciation to the Examiner for his courtesy in conducting an interview with Applicant's representative, Van Ernest of Greenblum & Bernstein, P.L.C., on August 11, 2006. The Examiner discussed the rejection under 35 U.S.C. § 101 and "practical applications" of the claims in more detail. No agreement was reached.

Upon entry of the present amendment, claim 16 will have been cancelled without prejudice and without disclaimer of the subject matter. Also, claims 1, 13 and 24 will have been amended to more clearly address a practical application of the claims. More particularly, claim 1 will have been amended to recite that whether to accept the prediction model is determined based on the comparison of the numerical value NTS with the numerical test statistic array; claim 13 will have been amended to substantially include the subject matter of canceled claim 16; and claim 24 will have been amended to recite that a hypothesis is accepted when the original data set is determined as not

arising by chance. Applicant respectfully submit that all pending claims are now in condition for allowance.

In the above-referenced Official Action, the Examiner rejected claims 1-34 under 35 U.S.C. § 101 for non-statutory subject matter, stating that the invention has not been limited to a substantial practical application. Applicant respectfully traverses this rejection, as discussed below, and further submits that the rejection is moot in view of the above amendments to the independent claims.

In addition, the Examiner rejected claims 1-7, 15-18, 20, 21, 26, 28-31 and 34 under 35 U.S.C. § 103(a) as being unpatentable over TATSUOKA (U.S. Patent No. 5,855,011) in view of ZHANG et at. (U.S. Patent No. 5,832,182). The Examiner rejected claims 8-11, 19, 22, 23, 27 and 33 under 35 U.S.C. § 103(a) as being unpatentable over TATSUOKA in view of ZHANG et at. and further in view of SHEN (U.S. Patent No. 6,041,788). The Examiner rejected claims 12-14, 24 and 25 under 35 U.S.C. § 103(a) as being unpatentable over TATSUOKA in view of ZHANG et al. and further in view of FAYYAD et al. (U.S. Patent No. 6,012,058). Applicant respectfully traverses these rejections at least for the following reasons.

The Examiner provided a number of bases regarding his rejection under 35 U.S.C. § 101. First, the Examiner stated that the claims have not been limited to a substantial practical application. Applicant submits that the rejection is now moot in view of the amendments to claims 1, 13, and 24, each of which now includes an added step for a practical application and a real world purpose. Applicant notes that the pending claims are at least analogous to the type of practical application to which the claims of the prior art relied upon by the Examiner (in the 35 U.S.C. § 103(a) rejections) are directed. For example, claim 1 in TATSUOKA is directed to "[a] method for classifying a test subject in one of a plurality of states in a domain ...," which, like the pending claims, may be generic, but is a practical, real world application nonetheless.

The Examiner also asserted that the recitations of the claims cannot be repeated and therefore lack concreteness, and are otherwise not reliable. Applicant respectfully submits that, if only one random number were used to examine a quantity of interest, then the Examiner's objection may be valid. However, because many random numbers are to be generated for each quantity of interest, the multiplicity of random data sets, each consisting of many random numbers (e.g., 1000 random data sets), provides a cumulative effect which is non-random. See, e.g., paras. [0064]; [0068]. In other words, Applicant submits that two analyses, each based on 1000 random data sets, would produce substantially the same results. Thus the results are substantially repeatable, predictable, reliable and hence concrete.

Similarly, the Examiner also asserted that comparing incoming data to random generated data has no function or real world purpose. However, as discussed above, Applicant respectfully asserts that large quantities of random numbers have the

cumulative effect of rendering the analysis repeatable, accurate, and precise. Thus comparing real data to randomly generated data does have a function and a real world purpose, e.g., to the extent a real world prediction model (in any discipline, such as medical research) can be accepted or rejected based on the determination of whether an event is random (i.e., happens by chance), as recited in claims 1, 13 and 24.

Lastly, the Examiner asserted that comparing real data to randomly generated data that is supposed to be of the same distribution is not reliable because one does not know the distributions due to the fact that it is randomly generated. However, the Examiner fails to take into consideration a basic distinction between data values and their distribution. The randomly generated data is generated according to predetermined, non-random constraints, and has a known distribution. The values are random but their distribution is not, and it is in fact possible to select the same distribution. Thus the recited analysis is reliable, as discussed above.

Accordingly, Applicant respectfully requests the Examiner to withdraw the rejections under 35 U.S.C. § 101.

Examiner's response:

There needs to be a real world function for the invention. As stated the invention is still in an abstract domain and provides no tangible purpose. As stated the invention is an exercise only. Examiner reminds the applicant that previous standards for 35 U.S.C. § 101 may not apply currently. Please see the interim guidelines for examination of patent applications for patent subject matter eligibility published November 22, 2005 in the official gazette.

Referring to using random data sets. Applicant admits that results are 'substantially repeatable'. 'Substantially repeatable' is not the same as exactly repeatable. 'Exactable repeatable' would mean predictable, reliable and concrete. 'Substantially repeatable' infers a percentage of error and thus not predictable, reliable and concrete. Examiner agrees that 'large quantities of random data' does have a cumulative effect and greatly increases the odds of repeatable outcomes but does not

make it 100 percent. "The randomly generated data is generated according to predetermined, non-random constraints, and has a known distribution" is in claim 2 and not claims 1, 13 and 24 and is addressed by Zhang. First Office Action applies.

7. In reference to the Applicant's argument:

With respect to the Examiner's rejections under 35 U.S.C. § 103(a), Applicant respectfully traverses numerous teachings asserted by the Examiner. With respect to independent claim 1, for example, the Examiner asserted that the NTS of the claim is equivalent to the SPS disclosed in TATSUOKA. However, the NTS is a numerical value obtained by application of a test statistic formula, whereas the SPS of TATSUOKA is a state probability set. A number cannot be a set, and so the two are not equivalent. Further, TATSUOKA discloses classification items, whereas claim 1 recites probability distributions. Since a probability distribution indicates the frequency of occurrence of different elements, such as items, probability distributions and items cannot be equivalent. Also, the Examiner asserted that "comparing" is equivalent to "test subject classified" of TATSUOK. However, "comparing," as used in claim 1, refers to the process of determining the relationship between a real number and the elements of an ordered set, while "test subject classified" refers to the assignment of a subject (i.e., of people) to a set. The two concepts are thus quite different.

Examiner's response:

Applicant is correct in stating that the two separate items confused the Examiner. The two items were the **NTS** and the **N**umerical **T**est **S**tatic array. The former is a value and the second is an array. Examiner has corrected the office action by mapping the NTS to the 'quality measure' of Tatsuoka. 'Test subject classified' is composed in part of a 'quality measure'. So comparing is done with both 'real numbers'. First Office Action applies.

8. In reference to the Applicant's argument:

The Examiner admitted that TATSUOKA does not teach creating a plurality of random data sets RDB(i) using randomly generated data, in which i is a positive integer. The Examiner therefore relied on ZHANG et al. to teach this feature. More particularly, the Examiner asserted that the "data sets" recitation in claim 1 is taught by "each intended cluster" disclosed in ZHANG et al. However, the term "data sets" refers to the entire set of data under consideration, while "each intended cluster" of ZHANG et al. refers to a collection of subsets (or clusters) of the data set under consideration. Hence the two concepts are not equivalent. Accordingly, Applicant respectfully requests withdrawal of the rejection of claim 1 based on the combination of at least TATSUOKA and ZHANG et al.

Examiner's response:

A collection of subsets is considered a set. Thus 'data sets' of applicant is equivalent to 'each intended cluster.' 'Each intended cluster' implies all clusters which is equivalent to datasets of applicant. First Office Action applies.

9. In reference to the Applicant's argument:

With respect to independent claim 13, Applicant amended the claim by further reciting that a determination is made, based on the comparison between the plurality of numerical values and the numerical value corresponding to the original data set, whether the original data set is characterized by at least one factor that is not based on chance (which is substantially the subject matter of canceled claim 16). First, this is a practical application satisfying 35 U.S.C. § 101, as discussed above. Second, the portion of TATSUOKA on which the Examiner relied (in rejecting claim 16) merely mentions "decision-theoretic rule," but does not teach or suggest comparing a fixed calculated value, derived from the original data set, to a set of calculated values derived from multiple random data sets (i.e., in order to determine whether a factor is based on chance). TATSUOKA does not make use of random data sets in this or any other respect. Accordingly, for at least this reason, in addition to the reasons set forth above with respect to claim 1, Applicant respectfully requests withdrawal of the rejection of claim 13 based on the combination of at least TATSUOKA and ZHANG et al.

Examiner's response:

Claim 13 incorporates Tatsuoka, Zhang and Fayyad. 'Decision theoretic rule' is a rule which in part is used for determining a classification. Classification is done by comparing 'quality measure' with each classification and deciding by comparison of the two values and is classified into the 'highest state.' Fayyad teaches random sets (**Fayyad,** C5:7-26, C1:30-37, C2:5-15) and is used in conjunction with Tatsuoka and Zhang. First Office Action applies.

10. In reference to the Applicant's argument:

With respect to independent claim 24, the Examiner admitted that the combination of TATSUOKA and ZHANG et al, do not teach each randomly generated data set having a second size, dimension and distribution relating to the original data set. The Examiner therefore relied on FAYYAD et at. to teach this feature. However, the "size" of the data sets in the claims refers to a measure of the entire original data set, while "M" of FAYYAD et at, relied upon by the Examiner, refers to the number of records in one cluster, which is merely a subset of the data. Further, contrary to the Examiner's assertion, FAYYAD et al. do not teach or suggest using input parameters such as size, dimension and distribution of data to have each randomly generated data set having a second size, dimension and distribution relating to the original data set. Although FAYYAD et at. discuss random sampling, it does not refer to even one random data set, let alone a plurality of random data sets. Further, FAYYAD et at. do not discuss a combination of size, dimension and distribution, whether relating to a data set, or relating to the comparison of data sets. Accordingly, for at least these reasons, in addition to the reasons set forth above with respect to claim 1, Applicant respectfully requests withdrawal of the rejection of claim 24 based on the combination of at least TATSUOKA, ZHANG et al. and FAYYAD et at.

The Examiner relied on SHEN only to teach creating a plurality of random data sets using randomly generated data according to a Monte Carlo technique. Therefore, SHEN does not overcome the deficiencies of the combination of TATSUOKA, ZHANG et at. and/or FAYYAD et at., discussed above.

With regard to claims 2-12, 14-23 and 25-34, Applicant asserts that they are allowable at least because they depend, directly or indirectly, from independent claims 1, 13 and 24, respectively, which Applicant submits have been shown to be allowable. Applicant further asserts that these dependent claims are allowable based on their respective recitations, as well. For example, claims 7 and 21 recite that the function of percentile indices is a linear combination of the non-empty set of percentile indices. The Examiner relied on TATSUOKA to teach this feature, but TATSUOKA does not mention percentile indices.

In view of the herein contained amendments and remarks, Applicant respectfully requests reconsideration and withdrawal of previously asserted rejections set forth in the Official Action of May 26, 2006, together with an indication of the allowability of all pending claims, in due course. Such action is respectfully requested and is believed to be appropriate and proper.

Any amendments to the claims in this Reply, which have not been specifically noted to overcome a rejection based upon the prior art, should be considered to have been made for a purpose unrelated to patentability, and no estoppel should be deemed to attach thereto.

Examiner's response:

Claim 24 incorporates Tatsuoka, Zhang and Fayyad. Fayyad illustrates generating data records of size 'M'. If there exists a second cluster of data records then there exists a second size of 'M'. Dimensions is illustrated by Fayyad with 'limited to small data sets with a small number of dimensions' which means a small vector length. 'Distribution' is illustrated by the 'model the probability distribution governing the data source' (**Fayyad,** C5:7-26, C1:30-37, C2:5-15) Tatsuoka illustrates a linear function using summation of percentile indices. (**Tatsuoka,** C6:51 through C7:4) The ratio of the two equations in (2) would be equivalent to the percentile indices. First Office Action applies.

Examination Considerations

11. The claims and only the claims form the metes and bounds of the invention. "Office personnel are to give the claims their broadest reasonable interpretation in light of the supporting disclosure. *In re Morris*, 127 F.3d 1048, 1054-55, 44USPQ2d 1023, 1027-28 (Fed. Cir. 1997). Limitations appearing in the specification but not recited in the claim are not read into the claim. *In re Prater*, 415 F.2d, 1393, 1404-05, 162 USPQ 541, 550-551 (CCPA 1969)" (MPEP p 2100-8, c 2, l 45-48; p 2100-9, c 1, l 1-4). The Examiner has the full latitude to interpret each claim in the broadest reasonable sense. Examiner will reference prior art using terminology familiar to one of ordinary skill in the art. Such an approach is broad in concept and can be either explicit or implicit in meaning.

12. Examiner's Notes are provided to assist the applicant to better understand the nature of the prior art, application of such prior art and, as appropriate, to further indicate other prior art that maybe applied in other office actions. Such comments are entirely consistent with the intent and sprit of compact prosecution. However, and unless otherwise stated, the Examiner's Notes are not prior art but link to prior art that one of ordinary skill in the art would find inherently appropriate.

13. Examiner's Opinion: Paragraphs 11 and 12 apply. The Examiner has full latitude to interpret each claim in the broadest reasonable sense.

Conclusion

14. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

15. Claims 1-15, 17-34 are rejected.

Correspondence Information

16. Any inquiry concerning this information or related to the subject disclosure should be directed to the Examiner Peter Coughlan, whose telephone number is (571) 272-5990. The Examiner can be reached on Monday through Friday from 7:15 a.m. to 3:45 p.m.

If attempts to reach the Examiner by telephone are unsuccessful, the Examiner's supervisor David Vincent can be reached at (571) 272-3687. Any response to this office action should be mailed to:

Commissioner of Patents and Trademarks,

Washington, D. C. 20231;

Hand delivered to:

Receptionist,

Customer Service Window,

Randolph Building,

401 Dulany Street,

Alexandria, Virginia 22313,

(located on the first floor of the south side of the Randolph Building);

or faxed to:

(571) 273-8300 (for formal communications intended for entry.)

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have any questions on access to Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll free).

Peter Coughlan

10/24/2006

JOSEPH P HIRL
PRIMARY EXAMINER
TECHNOLOGY CENTER 2100